Australian Government
Department of Defence
Defence Science and Technology Group

A Discussion of Dempster-Shafer Theory and its Application to Identification Fusion

E. El-Mahassni and K. White

National Security and ISR Division
Defence Science and Technology Group

DST-Group–TN–1443

ABSTRACT

This paper outlines some of the basics of Dempster-Shafer Theory, which is a mathematical theory for combining evidence from different sources to obtain a degree of belief in a proposition. In particular, different combination rules available within the context of Identification Fusion and the assumptions and implications of each of those rules are outlined and investigated. However, the belief function arising from combining evidence under Dempster-Shafer Theory is often insufficient to support decision-making, and a transformation from the belief function to a probability distribution is required. Several different transformations and illustrative examples implementing Dempster-Shafer Theory for Identification Fusion are provided. The results and some possible directions for future work are discussed.

RELEASE LIMITATION

Approved for Public Release

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
AUG 2015		

4. TITLE AND SUBTITLE	
A Discussion of Dempster-Shafer Theory and its Application to Identification Fusion	5a. CONTRACT NUMBER
	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Defence Science and Technology Group,PO Box 1500,Edinburgh, South Australia 5111, Australia, ,**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR’S ACRONYM(S)
	11. SPONSOR/MONITOR’S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT

This paper outlines some of the basics of Dempster-Shafer Theory, which is a mathematical theory for combining evidence from di erent sources to obtain a degree of belief in a proposition. In particular, di erent combination rules available within the context of Identi cation Fusion and the assumptions and implications of each of those rules are outlined and investigated. However, the belief function arising from combining evidence under Dempster-Shafer Theory is often insuf- cient to support decision-making, and a transformation from the belief function to a probability distribution is required. Several di erent transformations and illustrative examples implementing Dempster-Shafer Theory for Identi cation Fusion are provided. The results and some possible directions for future work are discussed.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**		**35**	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18

Published by

National Security and ISR Division
Defence Science and Technology Group
PO Box 1500
Edinburgh, South Australia 5111, Australia

Telephone: 1300 333 362
Facsimile: (08) 7389 6567

AR-016-371
August, 2015

APPROVED FOR PUBLIC RELEASE

A Discussion of Dempster-Shafer Theory and its Application to Identification Fusion

Executive Summary

Sensor Fusion is the combination of sensor data obtained from different sources to reduce uncertainty and/or obtain more complete or accurate information than that available from any individual source. In some cases, two or more sensors with complementary capabilities may be used together to obtain a target estimate that is better than when any individual sensor is used alone.

The scenario addressed by this paper is that of fusing the data from two or more sensors to support an operator's decision making ability about the identity of a target. This paper outlines some of the basics of Dempster-Shafer Theory (DST), which is a mathematical theory for combining evidence from different sources to obtain a degree of belief in a proposition. It overcomes the Bayesian constraint of needing precise fixed priors by assigning mass values which may or may not obey the classical probability axioms. This means DST is able to accommodate uncertain knowledge in a manner which proponents believe Bayesian probability cannot do, or at the very least is unable to in an obvious and straight-forward manner.

The aim of this paper is to describe the fusion of target identification data within the DST framework. Dempster's rule of combination is the most common DST approach for fusing data. Alternative rules of combination include Smets' Transferable Belief Model (TBM), Yager's Rule, Inagaki's Unified Combination Rule (UCR), and Dubois and Prade's rule. Each of the different rules of combination yields a distribution of belief or mass and the distribution may be different depending on the approach adopted. An examination of different rules of combination is presented and their underlying assumptions described. Simple examples are used to illustrate how the different rules of combination are implemented and to highlight the treatment of uncertainty in the sensor data. However, typically, evidence that is presented in the form of a distribution of belief does not directly enable decisions to be made. Alternative transformations from a belief to a probability distribution are described and illustrative examples presented.

THIS PAGE IS INTENTIONALLY BLANK

Contents

Glossary

ACR	Adaptive Combination Rule
AIS	Automatic Identification System
AU	Aggregation Uncertainty
bba	basic belief assignment
CWA	Closed-World Assumption
DST	Dempster-Shafer Theory
F	Friendly
FoD	Frame of Discernment
GPT	Generalized Pignistic Transformation
H	Hostile
ID	Identification
IRST	Infrared Search and Task
N	Neutral
OWA	Open-World Assumption
PCR	Proportion Conflict Resolution
PIC	Probabilistic Information Content
PiT	Pignistic Transformation
PlT	Plausibility Transformation
RF	Radio Frequency
RWR	Radar Warning Receiver
TBM	Transferable Belief Model
TVSU	Television Screening Unit

THIS PAGE IS INTENTIONALLY BLANK

1 Introduction

Sensor Fusion is the combination of sensor data obtained from different sources to reduce uncertainty and/or obtain more complete or accurate information than that available from any individual source. In some cases, two or more sensors with complementary capabilities may be used together to obtain a target estimate that is better than when any individual sensor is used alone. The scenario addressed by this paper is that of fusing the data from two or more sensors to support an operator's decision making ability about the identity of a target.

The most popular methods for combining information from multiple sensors are Bayesian in nature, including Bayesian networks and Kalman filters. Bayesian networks are models that represent variables and their related probabilities and are used to perform inference based on probability distributions provided by the sensor or other data sources [1]. Halpert extended target identification (ID) used Bayesian techniques for more than two targets [2], however Mortiss noted that this model does not work well for identifying neutral targets unless information on the nature and number of these targets is available prior to engagement [3]. Krieg showed how Bayesian Belief Networks could be used as a framework for linking target kinematics with attribute states [4], whilst a comprehensive tutorial in Bayesian Belief Networks was also given by the same author [5]. Maskell [6] addressed a well-known problem of handling conflicting information by setting appropriate values for the probability distributions. A key requirement for all Bayesian techniques is that the prior probability distributions must be specified in advance.

Dempster-Shafer Theory (DST) overcomes the Bayesian constraint of needing precise fixed priors by assigning mass values which may or may not obey the classical probability axioms. This means DST is able to accommodate uncertain knowledge in a manner which proponents believe Bayesian probability cannot, or at the very least is unable to do in an obvious and straight-forward manner.

As an example, suppose that $p(\alpha_1)$ and $p(\alpha_2)$ denote the probability of there being rain in a day or not respectively. Then clearly, $p(\alpha_1)+p(\alpha_2)=1$. However, if we let $m(\alpha_1)$ and $m(\alpha_2)$ depict the masses for there being rain or not, then under the DST framework $m(\alpha_1) + m(\alpha_2) + m(\alpha_1 \cup \alpha_2) = 1$, where $m(\alpha_1 \cup \alpha_2)$ denotes the uncertainty mass of not knowing whether it will rain or not. Note that if $m(\alpha_1 \cup \alpha_2) = 0$, then this is no more than the classical probability axiom, but this is not a requirement in DST. However, the computational complexity of DST can quickly escalate and while Bayesian probability is universally accepted, the same cannot be said of DST; partly due to it being relatively new (only a few decades old). Further, the concept of belief masses and what this means is still contested by some statisticians and mathematicians working in this area.

In effect, Dempster-Shafer Theory (DST) is another method for combining information from multiple sensors [7, 8, 9] and may be regarded as a collection of different theories with the same underlying uncertainty calculus and has its origins in the representation of and reasoning under partial probability distributions over a finite set of elements, called Θ. It encodes uncertainty by assigning normalised non-negative quantities (basic belief masses) to the elements of the power set of Θ (which are of order $2^{|\Theta|}$. Notwithstanding the fact that Random Set Theory has been proposed by Mahler as a unifying framework between different types of fusion rules [10, 11], the main benefits of DST are that it

specifically caters to the handling of ambiguity in systems that provides a more flexible means of representing ignorance than probability theory, and it is able to handle random variables with uncertain belief masses. Furthermore, DST has also recently been used with other techniques [12, 13, 14, 15, 16] for such purposes as rule-based inferencing and group decision-making among others.

In this paper, particular emphasis is placed on the fusion of information and sensor data from multiple sources which contains uncertainty. Bogler [17] detailed a use of Dempster's rule of combination using data from sensors such as a Radar Warning Receiver (RWR), a Television Sighting Unit (TVSU) and an Infra-Red Search and Track (IRST) sensor. DST has also been applied to multitarget tracking in clutter [18], whilst another approach which uses DST for joint tracking and classification/intent-detection can be found in [19]. Also recently, there has been some work done in trying to fuse soft data [20, 21, 22]. This is data that has a qualitative nature to it and is typically used in every day speech, for example "that car is near the city" or "that person is walking really fast". Apart from the fact that fixed priors must be known, or at best, it is sometimes argued that Bayesian methods can be too restrictive a fashion, Bogler also notes that Bayesian techniques do not explicity address the problem where the sensor data presents conflicting information. A recent survey of techniques that seek the address the problem of uncertainty from a theoretical and practical perspective is provided in [23]. More specifically, Hadzagic et al. [24] considered the relevance and reliability in assessing evidence for the implementation of DST to Electronic Support Measures. Koks and Challa [25] compared Bayesian fusion with DST, whilst Smets, as will be discussed later, also wrote extensively on the issue of comparing DST techniques with other methods for fusion, including specifically to the problem of target detection and ID.

Other recent interesting applications of DST involve the behaviour and decision-making aspect of travellers [26] and for the quantification and uncertainty propagation of AIS ship data [27]. The application considered in this paper is the problem of determining the ID, or identification, for a target of interest given data from multiple sensors. In this context, ID can be simply described as the task of determining or classifying the identity of unknown targets of interest. Depending on the specific requirements for the application, the target ID may refer to the unique platform (e.g. a specific ship such as *HMAS Anzac: FFH150*), the platform type (eg *Anzac Class Frigate*), the platform category (eg Frigate) and/or the platform allegiance (eg Friend). Data about the ID of a target may be received from heterogeneous sources including imaging, acoustic and RF sensors and even intelligence. The source data may relate to different levels of the target ID so that one sensor may be able to determine the allegiance while another may be able to determine platform type details.

The objective of sensor fusion is to estimate the states of targets given data from multiple sensors. For a given sensor, the target ID may be ambiguous, for example a sensor may not be able to resolve between several possible platform types. For many applications, the target state comprises kinematic and attribute components, that are estimated using Tracking and ID Fusion functions, respectively. A sequential process for sensor fusion whereby a tracking function is followed by ID fusion as shown in Figure 1 is considered. Data from each of the sensors is initially processed by a tracking function which performs data association to determine the sensor data that corresponds to the same target. Typically, a kinematic target state estimate is computed by the tracking function based on

kinematic components of the sensor data.

The focus of this paper is on ID fusion which has the goal of fusing ID data from multiple sensors to estimate the ID of targets. In the DST framework an ID Combination function processes the target ID evidence and obtains a belief or mass distribution. A Probability Transformation is then performed to determine a discrete probability distribution of the target ID. The output of Sensor Fusion is a track corresponding to each target with estimates of the kinematic state and of the target ID.

Sensor fusion supports decision making by processing uncertain sensor data to arrive at target state estimates. If a track on a target corresponds to a large freighter, then the decisions of an operator may be quite different than if the track corresponds to a small, fast vessel. Furthermore, a decision maker may be informed by the uncertainty in the target ID. For example, if it is equally likely that the target is a large freighter or a small, fast vessel then a decision maker might choose to make a conservative decision that is appropriate to their mission. Another course of action might be for the user to task sensors in an attempt to reduce the uncertainty in the estimate of the target ID.

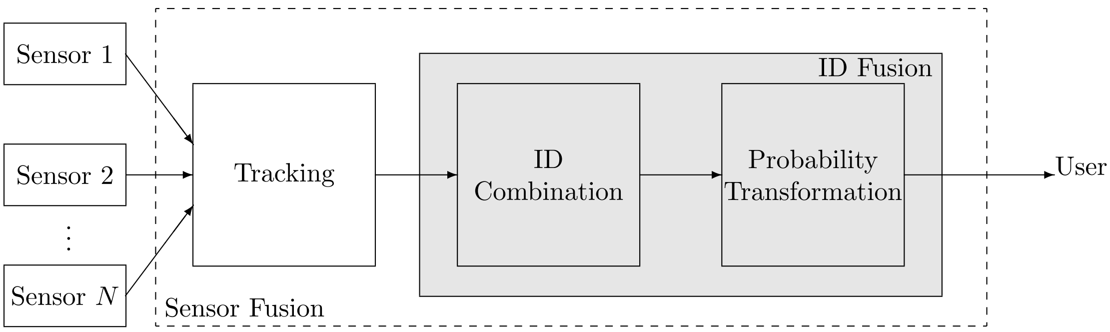

Figure 1: Functional flow for Sensor Fusion

The aim of this paper is to describe the fusion of target ID data within the DST framework. Dempster's rule of combination is the most common DST approach for fusing data. Alternative rules of combination include Smets' Transferable Belief Model (TBM), Yager's Rule, Inagaki's Unified Combination Rule (UCR), and Dubois and Prade's rule. Each of the different rules of combination yields a distribution of belief or mass and the distribution may be different depending on the approach adopted. An examination of different rules of combination is presented and their underlying assumptions described. Simple examples are used to illustrate how the different rules of combination are implemented and to highlight the treatment of uncertainty in the sensor data. However, typically, evidence that is presented in the form of a distribution of belief does not directly enable decisions to be made. Alternative transformations from a belief distribution to a probability distribution are described and illustrative examples presented.

2 Basic Concepts of Dempster-Shafer Theory

In this section, some of the basic concepts of DST are described. The set of possible target types that might be present is known as the frame of discernment (FoD). The construction of the FoD is outlined together with the belief function assignment, which is the designation of values or beliefs regarding target ID. Other issues like whether this FoD should be closed or open are also discussed.

2.1 Frame of Discernment (FoD)

At the heart of DST is the Frame of Discernment (FoD) which is typically denoted by Θ_D and is a model that describes the set of possible hypotheses.

Suppose that the FoD is

$$\Theta_D = \{\theta_1, \theta_2, \ldots, \theta_k\},$$

where the $\theta_i, 1 \leq i \leq k$, are mutually exclusive and contained in the set labelled Θ_D.

A Closed World Assumption (CWA) implies that Θ_D is exhaustive, so that at least one of the θ_i must be true. If it is known that θ_1 is not true, then the CWA implies that the conclusion must correspond to one of the other $k-1$ possibilities.

2.2 Expanding the Frame of Discernment (FoD)

It is possible to extend the FoD by combining it with another [28]. For instance, if we have a frame that describes the platform type so that

$$\Theta_1 = \{\theta_1, \theta_2, \theta_3\}.$$

then this FoD for platform type can be combined with another that describes the platform allegiance,

$$\Phi = \{\phi_1, \phi_2, \phi_3, \phi_4\},$$

A fully expanded FoD is obtained from the cross product:

$$\begin{aligned}\Theta_2^{(f)} &= \Theta_1 \times \Phi \\ &\quad \{(\theta_1, \phi_1), (\theta_1, \phi_2), \ldots, (\theta_3, \phi_4)\},\end{aligned}$$

which contains 12 elements. However, in many cases it may be known with certainty that not all elements in $\Theta_2^{(f)}$ will be present. For instance, the true FoD could be the subset of 5 elements:

$$\Theta_2 = \{(\theta_1, \phi_1), (\theta_1, \phi_2), (\theta_2, \phi_2), (\theta_2, \phi_3), (\theta_2, \phi_4)\}.$$

Further, there might be different levels of resolution so that

$$\Theta_3 = \{(\theta_1, \phi_1, \psi_1), (\theta_1, \phi_1, \psi_2), (\theta_1, \phi_2), (\theta_2, \phi_2), (\theta_2, \phi_3), (\theta_4, \phi_4, \psi_3), (\theta_4, \phi_4, \psi_4)\}.$$

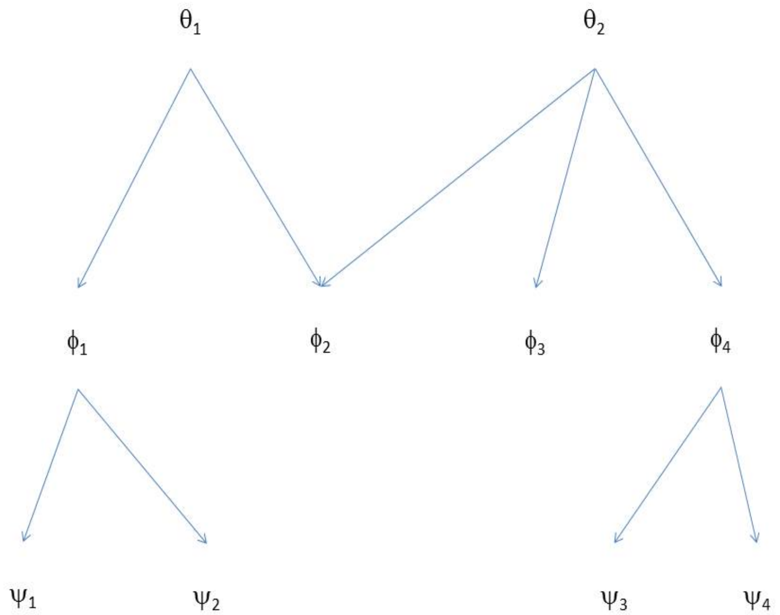

Figure 2: The FoD Θ_3 represented as an irregular tree

A depiction of this FoD in the form of an irregular tree is shown in Figure 2.

Turning to the problem of ID fusion, a key question is: How can masses from sensors which have different levels of resolution be fused? An answer to this question lies in forming appropriate subsets of the FoD that are applicable to Θ_2's FoD. For example, Θ_2 can be expressed as

$$\Theta_2 = \{((\theta_1, \phi_1, \psi_1) \cup (\theta_1, \phi_1, \psi_2)), (\theta_1, \phi_2), (\theta_2, \phi_2), (\theta_2, \phi_3), ((\theta_2, \phi_4, \psi_3) \cup (\theta_2, \phi_4, \psi_4))\}$$

so that if there is evidence for (θ_1, ϕ_1) then that evidence can be applied to the set union $(\theta_1, \phi_1, \psi_1) \cup (\theta_1, \phi_1, \psi_2)$, where $A \cup B$ means set of those elements which are either in A, B or in both.

2.3 Belief Function Assignment

In DST, uncertain evidence is encoded as a *belief function* or sometimes referred to as a *potential function* (this is at times distinctly used to distinguish it from Bayesian probability theory [29, 30]) which can be expressed in a variety of equivalent ways. There are several ways of encoding this uncertain evidence which include: mass, belief, plausibility and commonality. Denoting the function $[\phi]_m, [\phi]_b, [\phi]_{pl}, [\phi]_q$, etc can be used to indicate which representation of the potential is being used.

The most common means of encoding is via the *basic belief assignment* (bba) which is a function m defined on the power set of the frame of discernment Θ_D as follows:

$$\begin{aligned} m : \wp(\Theta_D) \rightarrow & \ [0,1] \\ X \mapsto & \ m(X) \end{aligned}$$

subject to the constraint that $\sum_{X \subseteq \Theta_D} m(X) = 1$, where $\sum_{k=1}^{n} a_k = a_1 + a_2 \ldots + a_n$. The power set of a set of elements is simply the set of all possible subsets of elements from that set. For instance, for a FoD $\Theta_1 = \{\theta_1, \theta_2, \theta_3\}$ consisting of $|\Theta_1| = 3$ elements, the power set is simply

$$\begin{aligned} \wp(\Theta_1) &= 2^{\Theta_1} \\ &= \{\{\emptyset\}, \{\theta_1\}, \{\theta_2\}, \{\theta_3\}, \{\theta_1, \theta_2\}, \{\theta_1, \theta_3\}, \{\theta_2, \theta_3\}, \{\theta_1, \theta_2, \theta_3\}\}. \end{aligned}$$

For a given set $X \subseteq \Theta_D$, the belief mass $m(X)$ represents the proportion of all relevant and available evidence that supports the claim that a particular element of Θ_D belongs to the set X, but to no particular subset of X. We remark that mathematically, m is akin to a probability distribution when the only non-zero belief masses are all assigned to subsets containing a single element, also known as a *singleton*. As already noted, no interpretation of the belief masses as probabilities is made in some interpretations of the DST.

To combine evidence from two independent belief functions represented as bbas m_1 and m_2, Dempster's rule of combination is typically used and the resulting belief function, denoted by $m_{1,2} = m_1 \oplus m_2$.

2.4 Belief, Plausibility and Commonality

Two other useful encodings in one-to-one correspondence with the basic belief assignment m are the *belief* and *plausibility* functions which are defined in the following manner:

$$\mathrm{bel}(X) = \sum_{\emptyset \neq Y \subseteq X} m(Y) \quad \text{and} \quad \mathrm{pl}(X) = \sum_{Y \cap X \neq \emptyset} m(Y).$$

For a given subset X, $\mathrm{bel}(X)$ quantifies the extent to which the evidence supports X, while $\mathrm{pl}(X)$ quantifies the extent to which the evidence does not contradict X.

Another encoding frequently used is commonality. It can be defined as the quantity of the agent's belief which may eventually be assigned to X. Mathematically, it can be represented as follows:

$$q(X) = \sum_{X \subseteq Y} m(Y),$$

where $A \subset B$ means the set that contains all those elements that A and B have in common. Its real value, however, lies in the fact that it can lead to much faster computation of masses (see [31]) due to the following property:

$$q_{12}(X) = q_1(X).q_2(X),$$

where $q_i, i = 1, 2$ are the commonality values assigned to X from two different sources and $q_{12}(X)$ is the fused commonality value of the q_is.

Note that DST can handle probabilities in the traditional manner, but this is not a requirement in order to use DST. However, a link can be made to standard probability measures by noting that belief and plausibility may be interpreted as lower and upper bounds on the probability, respectively. Note that in the remainder of this tutorial we adopt the conventional terminology where the term *belief function* is used to refer to the bba rather than bel(X).

2.5 Closed World Assumption (CWA)

Recall that the Closed World Assumption (CWA) implies that the FoD is exhaustive so that the solution must correspond to one of the elements in Θ_D. In other words, there is zero probability or chance that the true solution might lie outside of Θ_D.

Consider two situations where a CWA might be made. Firstly, the FoD may be known to be complete so that all possible solutions are listed. For target ID applications, such an assumption means that the sensors all have fully encoded knowledge of the possible platform types and are able to produce a belief mass distribution based on a library of possible platform types for subsequent ID Fusion. Secondly, the FoD may be known to be incomplete but is treated as complete so that the actual solution may lie outside of the FoD. In this situation, misidentification of a target may result from evidential reasoning applied to a FoD that is not exhaustive. Misidentification could lead to a target allegiance being declared hostile when in fact it should have been declared friend leading to fratricide or “friendly fire”. It is clearly important to be aware of the underlying assumption about the FoD.

2.6 Open World Assumption (OWA)

Lack of complete knowledge about the complete or entire solution space is known as having an Open World Assumption (OWA). Under the CWA a statement about a set element leads to inferences about the remaining set elements in the FoD. However, for an OWA, no such inference can be made. While the CWA may lead to inference errors if the frame of discernment is incomplete, the OWA places restrictions on the deductions that can be made from the available information.

Assuming a FoD Θ_D, the OWA assumption implies that if an inference is made about $\theta_1, \ldots, \theta_{k-1}$ then we cannot conclude anything about θ_k. The true FoD is actually

$$\Theta_F = \Theta_D \cup \Theta_E,$$

where Θ_E is another frame of unknown size so that $\Theta_D \cap \Theta_E = \emptyset$. That is, both frames have distinct elements and there is no “doubling-up” across each FoD.

3 Dempster's Rule of Combination

3.1 Definition

To combine evidence from two independent belief functions represented as bbas m_1 and m_2, Dempster's rule of combination is typically used and the resulting belief function, denoted by $m_{1,2} = m_1 \oplus m_2$, is given by the following direct sum:

$$(m_1 \oplus m_2)(X) = \frac{1}{(1-\kappa)} \sum_{Y \cap Z = X} m_1(Y) m_2(Z), \tag{1}$$

where the conflict mass κ is given by

$$\kappa = \sum_{Y \cap Z = \emptyset} m_1(Y) m_2(Z). \tag{2}$$

In Dempster's rule of combination, the empty set mass is zero, i.e. $m(\emptyset) = 0$ [32]. It should be noted that Dempster's rule is commutative, i.e. $m_1 \oplus m_2 = m_2 \oplus m_1$ and associative, i.e. $(m_1 \oplus m_2) \oplus m_3 = m_1 \oplus (m_2 \oplus m_3)$, when the basic belief masses are compatible, i.e. are not in complete conflict.

3.2 Conflict

A conflict between two beliefs in DST is where one source strongly supports one hypothesis and the other strongly supports another hypothesis, so that the two hypotheses are not compatible. The conflict mass κ quantifies the disparity of information between two sources and is non-zero whenever there is an incompatibility in the hypotheses corresponding to each source [33]. The closer that κ is to 1, the more we should reconsider how the evidence influences the beliefs [34]. If fusing two pieces of information (so that one has a belief in X and the other in $\bar{X}$) result in $\kappa = 1$ then it is clear that this fusion does not justify any support for either option (viz either X or $\bar{X}$).

Zadeh [35] showed that normalization in Dempster's rule would apparently resolve conflict but might result in misleading final values, whilst [36] describes a novel technique for discounting evidence in proportion to the degree of conflict it provides. An example was presented where two medical doctors (h_1 and h_2) inspected a patient and provided diagnoses for meningitis (M), concussion (C) and tumor (T). Their advice is listed in Table 1.

Table 1: Medical Diagnoses of Two Doctors

Medical	Diagnosis		
Doctor	M	C	T
h_1	0.99	0.00	0.01
h_2	0.00	0.99	0.01

In this example when all masses are combined according to Dempster's rule, we have the counter-intuitive result that the fused masses give a value of 1 for tumor. This result

arises because the only common non-zero masses are those that correspond to tumor. Other difficulties with DST have been discussed by Dezert et.al. [37, 38, 39]. Bayesian fusion techniques give a similar result to DST approaches when prior information is not carefully defined [6].

4 Alternative Rules of Combination

Dempster's Rule of Combination is the most common DST approach for fusing information from multiple sources. In this section we describe the following approaches in which evidence from multiple sources is fused:

- Smets' Transferable Belief Model (TBM)
- Yager's rule
- Inagaki's Unified Combination Rule (UCR)
- Dubois and Prade's rule

Note that these are probably the most well-known transformations and each make subtle assumptions about the nature of the FoD as will be discussed shortly. Furthermore, a different approach for treating evidence implies that there is a different approach for the management of conflicting information in the source data. The management of conflict is described for each of the different rules of combination.

4.1 Smets' Transferable Belief Model

The most well-known rule for explicitly accounting for conflict is the Transferable Belief Model (TBM) [40]. In this case, the combination rule is given by

$$m^s(X) = (m_1 \oplus m_2)(X) = \sum_{Y \cap Z = X} m_1(Y) m_2(Z), \tag{3}$$

where the mass for the empty set $m(\emptyset) \geq 0$ in contrast to Dempster's rule of combination and many other rules of combination where $m(\emptyset) = 0$.

The effect of normalisation with the conflict mass κ in Dempster's rule of combination is that a closed world is assumed. However, by allowing the possibility of the empty set mass to be non-zero in Smets' TBM, this implies that an open world is assumed. The TBM is commutative and associative, just like Dempster's rule of combination. It is worth noting that Smets has co-authored a number of papers regarding the application of TBM to target detection [41, 42, 43, 44, 45, 46, 47, 48].

4.2 Yager's Rule

Yager's rule of combination is similar to Shafer and Smets' TBM but it differs in the treatment of mass that is assigned to the empty set [49]. Yager defines the *ground probability assignment* associated with X as

$$r(X) = \sum_{Y \cap Z = X} m_1(Y)m_2(Z).$$

The bba arising from Yager's rule of combination is given by

$$\begin{aligned} m^y(X) &= r(X) &\quad (4)\\ m^y(\Theta_D) &= r(\Theta_D) + r(\emptyset) &\quad (5)\end{aligned}$$

so that any conflict $\kappa = r(\emptyset)$ is assigned to the FoD Θ_D. That is, Yager's rule implies a CWA because the possible solutions are contained in the FoD Θ_D.

We can relate Yager's rule to Dempster's rule in the following manner

$$\begin{aligned} m^d(\emptyset) &= 0\\ m^d(\Theta_D) &= \frac{r(\Theta_D)}{1-r(\emptyset)}\\ m^d(X) &= \frac{r(X)}{1-r(\emptyset)}.\end{aligned}$$

Note that Yager's rule is commutative but not associative. Rather, Yager's rule is considered to have a "quasi-associative" property, which means that the operator can be broken down into associative suboperations [32]. Yager justified this property on the basis that sometimes the order of operation does matter; in particular when new information is available and this is incorporated into an already combined structure [49].

4.3 Inagaki's Unified Combination Rule

Inagaki's UCR is a parametrized class of combination operations that incorporates both Dempster' rule and Yager's rule. Specifically, Inagaki [50] argues that every combination rule can be expressed in the form:

$$m^i(X) = r(X) + f(X)r(\emptyset),$$

where $X \neq \emptyset$, with $m^i(\emptyset) = 0$, and

$$\sum_{X \subset \wp(\Theta_D), X \neq \emptyset} f(X) = 1, \;\; f(X) \geq 0.$$

From these equations Inagaki's unified combination rule can be derived so that:

$$\begin{aligned} m^i_\alpha(X) &= [1+\alpha r(\emptyset)]q(X), \text{ where } X \neq \wp(\Theta_D), \emptyset &\quad (6)\\ m^i_\alpha(\Theta_D) &= [1+\alpha q(\emptyset)]q(\Theta_D) + [1+\alpha r(\emptyset) - \alpha]r(\emptyset) &\quad (7)\\ & 0 \leq \alpha \leq \frac{1}{1-r(\emptyset)-r(\Theta_D)} &\quad (8)\end{aligned}$$

If $\alpha = 0$ then the results are equivalent to Yager's rule and if $\alpha = 1/[1-r(\emptyset)]$, then the masses are equivalent to those obtained using Dempster's rule. The Inagaki's UCR is commutative, but quasi-associative.

4.4 Dubois and Prade's Rule

Dubois and Prade [51] took a set-theoretic approach in arriving at their rule

$$(m_1 \oplus m_2)^{dp}(X) = \sum_{Y \cup Z = X} m_1(Y) m_2(Z). \tag{9}$$

The use of the set union in Dubois and Prade's rule avoids consideration of the conflict mass and so there is no requirement for normalisation. The Dubois and Prade rule of combination is commutative and associative.

4.5 Other Rules of Combination

There are other rules of combination which have been developed within the DST framework such as Zhang's Center Combination Rule (CCR) [52], Mixing or Averaging [53], the Adaptive Combination Rule (ACR) [56], Convolutive X-Averaging [53], the Cumulatative rule [54], the Cautious rule [55], the collection of Proportion Conflict Resolution (PCR) rules [37, 56] and Zhang's Center Combination Rule (CCR) is the same as Dempster's for some cases and scaling problems may arise in some instances ([32]). The Mixing or Averaging, the Cumulative and the ACRs rely on weights being assigned to the mass distributions, whilst the Convolutive X-Averaging deals with information given as intervals. The Cautious rule assumes dependency among the evidence sources which is not considered here, while the PCR set of rules are 5 rules proposed use a variant or variants of the conjunctive rule to redistribute the conflicting mass on non-empty sets according to integrity constraints. Finally, Dubois and Prade [57] proposed another rule which combines both their original disjunctive rule with Smets' TBM. These are not described further in this paper but thorough reviews of various DST rules of combination have been provided elsewhere such as in [56], [58] and especially [32]. Lastly, most of these rules are also more thoroughly reviewed by Nguyen and Docking ([59]).

5 Probability Transformations

The bba distribution of the DST framework does not readily support decisions to be undertaken. Transformations are required to convert from the bba function to a probability distribution, however other approaches to decision-making not involving transformations have also been proposed [12, 15, 16]. This ID fusion approach where ID combination is followed by a probability transformation corresponds to two levels of reasoning within the DST framework. At the *credal* level, evidence is represented and manipulated using rules of combination, while at the *pignistic* level, evidence is manipulated into a probability distribution to support decision-making.

In a similar manner to the rules of combination, there are several probability transformations within the DST framework. Four different probability transformations are described here:

- Aggregate Uncertainty Approach (AU)

- Plausibility Transformation (PlT)
- Pignistic Transformation (PiT)
- Generalised Pignistic Transformation (GPT)

Each of the probability transformations result in a distribution of probability $P(X)$ over singleton elements, where $|X| = 1$, of the FoD Θ_D. A decision may then be made by, for example, declaring the target type that is most highly supported by the amassed evidence to be X_k, where $X_k = \arg_{\max(i=1,\ldots,k)} P(X_i)$.

The probability transformations listed above are among the most popular and they obey the *minimal criterion* which essentially states that the result of applying the transform to the vacuous belief function should be the uniform probability distribution [60]. A vacuous belief function is where the ignorance mass, or the mass over the entire FoD is 1. Other probability transformations not considered here include those of Sudano [61] and Cuzzolin [62].

5.1 Aggregate Uncertainty Approach

The Aggregate Uncertainty (AU) approach method distributes the masses of non-singleton elements to the constituent singletons and then normalises to arrive at a probability mass [63]. For instance, the mass $m(Y \cup Z)$ is distributed to $m(Y)$ and $m(Z)$, where $|Y| = |Z| = 1$. The AU method selects the proportion of $m(Y \cup Z)$ distributed to each of the singletons in such a way as to minimise the entropy of the resulting probability distribution [8]. Again, for a binary frame (i.e. Two elements), this process is straightforward: if $m(X) \geq 0.5$, the transform yields $P_1(X) = m(X)$ and $P_1(Y) = m(Y) + m(X \cup Y)$, otherwise $P_1(X) = P_1(Y) = 0.5$.

5.2 Plausibility Transformation

Also known as the Bayesian approximation method,the Plausibility Transformation (PlT) has been argued that this transformation is the most consistent technique for translating Dempster's rule of combination to the Bayesian probability domain [64]. For a binary frame, if $X \in \Theta_D$, where $a \in A$ indicates that the element a belongs to the set A, the probabilities given by this transformation become

$$P_2(X) = K^{-1}\mathrm{pl}(X),$$

where $K = \sum_{Y \in \Theta_D} \mathrm{pl}(Y)$.

5.3 Pignistic Transformation

The Pignistic Transformation (PiT) [40] moves the belief mass from the union elements of the power set and distributes it equally amongst the corresponding singleton members.

For an element $X \in \Theta_D$, the PiT gives

$$P_3(X) = \sum_{\emptyset \neq Y \subseteq \wp(\Theta_D):X \in Y} \frac{m(Y)}{(1-\kappa)|Y|}, \tag{10}$$

where κ is defined as before.

An OWA is made under TBM. However, once the pignistic transform is required to make a decision it is then really apportioning masses to the singletons as though a CWA is being made.

5.4 Generalized Pignistic Transformation

The Generalized Pignistic Transformation (GPT) aims to maximise the Probabilistic Information Content (PIC). Here it suffices to say that maximization of the PIC is equivalent to minimization of Shannon's entropy [37]. For an element $X \in \Theta_D$, the transformation is given by

$$P_4(X) = m(X) + (m(X) + \varepsilon) \sum_{\substack{Y \in \wp(\Theta_D) \\ X \subset Y, |Y| \geq 2}} \frac{m(Y)}{\sum_{\substack{Z \in \wp(\Theta_D) \\ Z \subset Y, |Z|=1}} m(Z) + \varepsilon |Y|}, \tag{11}$$

for some small arbitrary $\varepsilon \geq 0$.

6 Example Sensor Data

In this section we describe several different cases that are used to illustrate the application of DST to target ID.

First, we consider a FoD that describes allegiance so that

$$\Theta_A = \{F, H, N\},$$

where F, H and N represent friendly, hostile or neutral allegiance, respectively. Further, for each allegiance type, there are certain associated platform types. This can be represented by another FoD:

$$\Theta_P = \{A, B, C, D, E\}.$$

Next, combining these frames of discernment a fully expanded FoD with 15 platform types (that include allegiance) is obtained. These are:

$$\Theta_D^{(f)} = \{FA, FB, \ldots, NE\},$$

where the first letter represents the allegiance and the second indicates the platform type. In addition, if it is known with certainty that the true target ID does not correspond to certain elements within the FoD, then these may be pruned from further consideration so that the FoD is reduced to a smaller set of elements, as in the example of Θ_D below:

$$\Theta_D = \{FA, FC, FD, HB, HC, NA, NC, NE\}.$$

Figure 3 shows the FoD for target ID depicted as a tree of two layers with allegiance and platform type, respectively. Reducing the FoD in this manner is one option to reduce the computation complexity. However, from a modelling perspective, preserving the original frame with 15 elements might be preferable in case there is evidence that could be used to break the deadlock. Also, from the point of view of maintaining the underlying database or ontology, this might provide a cleaner solution.

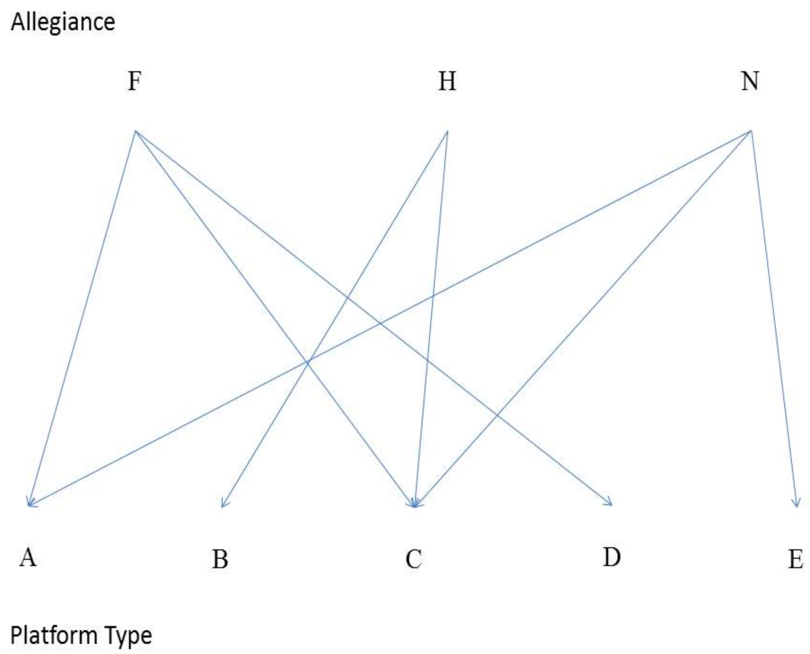

Figure 3: A tree representing the FoD for the example listed above

As an aside, on occasions, platform types may be combined so that FAC may be used in the instance where sensors may not be able to discern between platform types FA and FC. In a similar manner, other classification labels could be added to include target environment such as air, surface and land.

We assume that there is a single target of interest and that sensor data relating to the target ID is to be processed. Five different cases are considered in which uncertain data about the target ID is received from two or more sensors. Table 2 shows the example sensor input data for ID fusion. There are two sensors under consideration for cases 1 to 4 while for case 5 there are three sensors. The sensor input data is represented as a discrete mass distribution over the FoD where the sum of the masses for each sensor is unity. A blank entry in Table 2 implies that the corresponding mass is zero.

Table 2: Example sensor input data for ID fusion

Case	Sensor	Target ID								
		FA	HB	NA	NC	$FA \cup NA$	F	H	N	Θ_D
1	$S1$	0.7	0.2	0.1						
	$S2$	0.6	0.2		0.2					
2	$S1$	0.7	0.2	0.1						
	$S2$		0.4			0.6				
3	$S1$	0.7	0.2	0.1						
	$S2$						0.6	0.2	0.2	
4	$S1$	0.7	0.2							0.1
	$S2$					0.6				0.4
5	$S1$	0.7	0.2	0.1						
	$S2$	0.6	0.2		0.2					
	$S3$			0.6						0.4

In case 1, the two sensors provide belief masses for singleton elements. Sensor $S1$ provides non-zero masses for the elements FA, HB and NA while the data for sensor $S2$ differs in the values for FA, HB and the mass for NA is zero while the mass for NC is non-zero.

In case 2, sensor $S1$ provides the same mass distribution as for case 1. However sensor $S2$ provides non-zero mass for element HB and for the union $(FA \cup NA)$. In other words, sensor $S2$ is able to discern that the platform type is either HB or one of FA or NA, without being more specific about the latter two types.

Some sensors may only be able to discern data at the allegiance level. In case 3 sensor $S1$ provides the same mass distribution as for cases 1 and 2, but sensor $S2$ provides data at the allegiance level only so the mass distribution is for the elements of F, H and N. The allegiance elements may be expanded in terms of their constituent platform types, ie. $F = \{FA \cup FC \cup FD\}$, $H = \{HB \cup HC\}$ and $N = \{NA \cup NC \cup NE\}$.

In case 4 each sensor attributes some of the mass to elements of the FoD and the remaining mass is attributed to the FoD Θ_D. This is a means of handling missing information.

In case 5 there are three sensors. Sensors $S1$ and $S2$ provide mass distributions for singletons while sensor $S3$ provides mass for NA and Θ_D.

7 Results

In this section we present the results of applying ID fusion to the example sensor data described in the preceding section. We consider ID combination first where the DST rules of combination are applied to obtain a combined bba from multiple bbas provided by each sensor. Then, a probability transformation is applied to transform the combined bba into a probability distribution.

7.1 ID Combination results

For the purposes of illustrating how the ID combination rules are applied, we adopt a tabular representation similar to that of Sentz et al [32], as shown in Table 3. The following steps show how Dempster's rule of combination in equations (1) and (2) may be implemented for the example data corresponding to case 2. The fused mass is assigned to the intersection of the elements of both masses being fused. The second line in each entry in Table 3 indicates the resulting element to which the fused mass is assigned.

1. Compute the combined mass for an element as the product of the masses from the corresponding row and column. For instance,

$$m_1(FA)m_2(HB) \quad = \quad 0.28,$$

   contributes to κ since $\{FA\} \cap \{HB\} = \emptyset$. Or

$$m_1(FA)m_2(FA \cup NA) \quad = \quad 0.42,$$

   contributes to the $m_1 \oplus m_2(FA)$ since $\{FA\} \cap \{FA \cup NA\} = \{FA\}$.

2. Conflicting evidence arises when the intersection is empty, eg. $FA \cap HB = \emptyset$. The product of the masses from the corresponding row and column contributes to the summation for the conflict κ in equation 2.

3. Sum the masses for the conflict:

$$\begin{aligned}\kappa &= \sum_{Y \cap Z = \emptyset} m_1(Y)m_2(Z) \\ &= m_1(FA)m_2(HB) + m_1(NA)m_2(HB) + m_1(HB)m_2(FA \cup NA) \\ &= 0.28 + 0.04 + 0.12 = 0.44\end{aligned}$$

4. Where the intersection is not empty, this represents supporting evidence for the set intersection. The product of the masses from the corresponding row and column contributes to the summation for the mass of the relevant set in equation 1, eg. $\{FA\} \cap \{FA \cup NA\} = \{FA\}$.

5. Sum the masses for each element, eg:

$$\begin{aligned}(m_1 \oplus m_2)(FA) &= \frac{1}{(1-\kappa)} \sum_{Y \cap Z = FA} m_1(Y)m_2(Z), \\ &= \frac{1}{(1-0.44)} m_1(FA)m_2(FA \cup NA) \\ &= \frac{0.42}{0.56} = 0.75\end{aligned}$$

The results for ID combination using the different DST rules of combination for each of the 5 cases are presented in Table 4 with the notation $m^{(Y,Z)} = m(Y \cup Z)$. The value of the parameter α in Inagaki's Unified Combination Rule was selected to be $1/[2(1 - r(\emptyset))]$, which is the halfway point between Yager's rule and Dempster's rule.

Table 3: Example of Dempster's rule of combination applied to case 2

			Sensor 2		
			ID	HB	$FA \cup NA$
			m_1	0.4	0.6
Sensor 1	ID	m_2			
	FA	0.7		$m_1(FA)m_2(HB) = 0.28$, $FA \cap HB = \emptyset$	$m_1(FA)m_2(FA \cup NA) = 0.42$, $FA \cap (FA \cup NA) = FA$
	HB	0.2		$m_1(HB)m_2(HB) = 0.08$, $HB \cap HB = HB$	$m_1(HB)m_2(FA \cup NA) = 0.12$, $HB \cap (FA \cup NA) = \emptyset$
	NA	0.1		$m_1(NA)m_2(HB) = 0.04$, $NA \cap HB = \emptyset$	$m_1(NA)m_2(FA \cup NA) = 0.06$, $NA \cap (FA \cup NA) = NA$

It can be seen in Table 4 that the belief masses arising from each of the ID combination rules sum to one, as required. A coarse inspection reveals that Dubois and Prade's disjunctive rule typically yields a belief mass distribution for a greater number of elements than the other rules of combination. On the other hand, Dempster's rule usually produces non-zero belief masses for the least number of set elements because it generally does not assign mass to the empty set, $\emptyset$ or to the FoD Θ_D, like Dubois and Prade's rule. The exception to this is in case 4 when Θ_D appears in the sensor input data.

7.2 Probability Transformation results

Probability transformations are applied to compute probability values from the belief mass distributions obtained from the ID combination function. The belief masses obtained by applying the Dubois and Prade rule of combination to case 2 are used to illustrate the different probability transformations. The reason for selecting the results of the Dubois and Prade rule of combination is simply that the corresponding bba for case 2 has nonzero masses for elements with cardinality of one, two and three. In Table 5, P_1, P_2, P_3 and P_4 denote the AU, the PiT, the PlT and the GPT respectively as applied to the bba arising from the Dubois and Prade rule of combination for case 2. For the GPT the value of ϵ was set to 0.001.

8 Discussion

Several interesting features of ID fusion using DST can be observed from the results tabulated in Tables 4 and 5. A comparison of the results in Table 4 for Smets' and Yager's rules shows that the belief mass distribution are usually similar in that conflict is distributed to the empty set mass in the former and to the ignorance or universal mass in the latter. Dubois and Prade's rule yields similar results to Yager's rule. This is not

Case	Rule of Combination				
	Dempster	Smets	Yager	Inagaki	Dubois & Prade
1	$0.91^{(FA)}$ $0.09^{(HB)}$	$0.42^{(FA)}$ $0.04^{(HB)}$ $0.54^{(\emptyset)}$	$0.42^{(FA)}$ $0.04^{(HB)}$ $0.54^{(\Theta_D)}$	$0.67^{(FA)}$ $0.06^{(HB)}$ $0.27^{(\Theta_D)}$	$0.42^{(FA)}$ $0.26^{(FA,HB)}$ $0.04^{(HB)}$ $0.06^{(FA,NA)}$ $0.02^{(HB,NA)}$ $0.04^{(HB,NC)}$ $0.04^{(NA,NC)}$ $0.14^{(FA,NC)}$
2	$0.75^{(FA)}$ $0.14^{(HB)}$ $0.11^{(NA)}$	$0.42^{(FA)}$ $0.08^{(HB)}$ $0.06^{(NA)}$ $0.44^{(\emptyset)}$	$0.42^{(FA)}$ $0.08^{(HB)}$ $0.06^{(NA)}$ $0.44^{(\Theta_D)}$	$0.59^{(FA)}$ $0.11^{(HB)}$ $0.08^{(NA)}$ $0.22^{(\Theta_D)}$	$0.04^{(HB,NA)}$ $0.28^{(FA,HB)}$ $0.08^{(HB)}$ $0.48^{(FA,NA)}$ $0.12^{(FA,HB,NA)}$
3	$0.88^{(FA)}$ $0.08^{(HB)}$ $0.04^{(NA)}$	$0.42^{(FA)}$ $0.04^{(HB)}$ $0.02^{(NA)}$ $0.52^{(\emptyset)}$	$0.42^{(FA)}$ $0.04^{(HB)}$ $0.02^{(NA)}$ $0.52^{(\Theta_D)}$	$0.65^{(FA)}$ $0.06^{(HB)}$ $0.03^{(NA)}$ $0.26^{(\Theta_D)}$	$0.42^{(FA,FC,FD)}$ $0.04^{(HB,HC)}$ $0.06^{(NA,FA,FC,FD)}$ $0.04^{(HB,NA,NC,NE)}$ $0.02^{(NA,HB,HC)}$ $0.12^{(HB,FA,FC,FD)}$ $0.14^{(FA,NA,NC,NE)}$ $0.14^{(FA,HB,HC)}$ $0.02^{(NA,NC,NE)}$
4	$0.80^{(FA)}$ $0.09^{(HB)}$ $0.07^{(FA,NA)}$ $0.05^{(\Theta_D)}$	$0.70^{(FA)}$ $0.08^{(HB)}$ $0.06^{(FA,NA)}$ $0.12^{(\emptyset)}$ $0.04^{(\Theta_D)}$	$0.70^{(FA)}$ $0.08^{(HB)}$ $0.06^{(FA,NA)}$ $0.16^{(\Theta_D)}$	$0.75^{(FA)}$ $0.09^{(HB)}$ $0.06^{(FA\cup NA)}$ $0.10^{(\Theta_D)}$	$0.42^{(FA,NA)}$ $0.12^{(FA,HB,NA)}$ $0.46^{(\Theta_D)}$
5	$0.91^{(FA)}$ $0.09^{(HB)}$	$0.17^{(FA)}$ $0.02^{(HB)}$ $0.82^{(\emptyset)}$	$0.17^{(FA)}$ $0.02^{(HB)}$ $0.32^{(NA)}$ $0.49^{(\Theta_D)}$	$0.37^{(FA)}$ $0.04^{(HB)}$ $0.23^{(NA)}$ $0.37^{(\Theta_D)}$	$0.29^{(FA,NA)}$ $0.16^{(FA,HB,NA)}$ $0.08^{(FA,NA,NE)}$ $0.02^{(HB,NA,NE)}$ $0.01^{(NA,NE)}$ $0.04^{(HB,NA)}$ $0.40^{(\Theta_D)}$

Table 4: Belief mass distributions obtained using different DST combination rules for five cases of sensor data input.

surprising given that they both assume a closed world. However, whilst Yager's rule assigns any conflict to the universal mass, Dubois and Prade's rule assigns its conflict to the union of the elements of the fused masses.

Table 5: Results using the different probability transformations applied to the belief mass distribution obtained from Dubois and Prade's rule of combination for case 2.

ID	Probability Transformation			
	Aggregate Uncertainty, $P1$	Plausibility Transformation, P_2	Pignistic Transformation, P_3	Generalized Pignistic Transformation, P_4
FA	0.33	0.43	0.42	0.25
HB	0.33	0.26	0.28	0.51
NA	0.33	0.31	0.3	0.24

But, as can be observed from the last example, this method yields the trivial result when the output from one source is completely ignorant. It is interesting to note for all the rules, except Dubois and Prade's, if even one sensor does not contain a value for an element of the power set (say h_i), but all the other sensors or sources do, then when fused no mass will be allocated to h_i. This is in effect what occurred with Zadeh's example; even though a doctor gave a large confidence value for a particular diagnosis, a value of zero from the other doctor effectively "overpowers" that diagnosis. Indeed, a value of zero or one, is very profound and is something to bear in mind when using the aforementioned rules. Hence, it is better to assign at least some amount of non-zero mass to the FoD to offset this phenomenon. This is analogous to the rule of practice of using a non-zero likelihood function in Bayesian reasoning. The other interesting example is case 5 using Dubois and Prade's rule. Here, we see that even though sensor 3 does not in effect add or take away any information, fusing it with the other 2 sensors renders maximum ignorance. In this case, the ignorance of this sensor overpowers any other information provided by the other two.

The results for the probability transformations in Table 5 demonstrate that the AU rule and the GPT have quite different outcomes. The AU rule aims to maximise the entropy, whilst the GPT aims to minimise it. In terms of uncertainty, the AU transformation is the most pessimistic, while the GPT is the most optimistic. The PiT and PlT sit somewhere in between. Advocates of the PiT claim that it takes the most prudent betting criterion. Advocates of the PlT claim that it most closely resembles Bayesian probability.

Typically the platform type with the largest probability might be reported to a user, however for AU this is not possible. Note that the platform type with the highest probability for P_4 differs from that of P_2 and P_3. The occurrence of different decisions suggests that careful selection of the probability transform is needed. In many target ID applications, exceedance of a specified threshold is required before any decision on the platform type is made. Further, if a decision is made so that the declared platform type is that which is given by the maximum probability then just from this example above, this will vary according to the probability transformation. For the AU, all platform types are equally probable and it would be interesting to find out if this rather dramatic result is due to the AU transformation itself or rather as a result of Dubois and Prade's transformation. For the PlT and the PiT, FA is the most likely platform type, whilst for the GPT, HB is the

most likely platform type. As an example, if a threshold of 0.5 was applied for a decision, then only the GPT would yield a decision for the platform type, even though out of all the transformations, this is the only one that would conclude that HB was the most likely target.

9 Conclusion

In this paper, an overview of DST and its application to the problem of ID fusion using data from multiple sensors have been presented. Several examples have been provided to demonstrate various rules of combination and probability transformations in the DST framework. Underlying assumptions for open and closed worlds are paramount in guiding the selection of the most appropiate ID combination rule. With this in mind, further work will concentrate on the significance of assuming an open versus a closed world assumption with regards to target identification and classification and, by implication, the different ways in which conflicting reports are resolved at a decision-making level. In addition, an understanding on when perhaps one rule might be preferred in this type of application will also be the subject of future study, as is the investigation of how evidence received at different time instants may be updated in the DST framework. Future work will focus on comparing DST to Bayesian fusion for target ID.

References

1. Niedermeyer, D. *An Introduction to Bayesian Networks and Their Contemporary Applications*, Innovations in Bayesian Networks, Studies in Computational Intelligence (SCI), Holmes, D. and Jain, L. (Eds) **156**, 2008, 117-130, 2008.

2. Haspert, J. K. *Optimum ID Sensor Fusion for Multiple Target Types*, IDA Document D-2451, 2000.

3. Mortiss, G. *A Review of the Haspert Model for Target Identification*, DSTO Technical Note, DSTO-TN-0398, 2001.

4. Krieg, M. *Kinematic and Attribute Fusion Using a Bayesian Belief Network Framework*, DSTO Research Report, DSTO-RR-0315, 2006.

5. Krieg, M. *A Tutorial on Bayesian Belief Networks*, DSTO Technical Note, DSTO-TN-0403, 2001.

6. Maskell, S. *A Bayesian Approach to Fusing Uncertain, Imprecise and Conflicting Information*, Inf. Fusion, **9**, 259-277, 2008.

7. Dempster, A. P. *A Generalisation of Bayesian Inference*, J. Royal Stat. Soc., Series B **30**, 205-247, 1968.

8. Shafer, G. *A Mathematical Theory of Evidence*, Princeton University Press, 1976.

9. Shafer, G. *Dempster-Shafer Theory*, Encyclopedia of Artificial Intelligence 2nd Ed., Shapiro, S. C. (Ed), Wiley, 330-331, 1992.

10. Mahler, R. *Representing Rules as Random Sets, I: Statistical Correlations Between Rules*, Inf. Sciences, **88**,47-68, 1996.

11. Mahler, R. *Representing Rules as Random Sets, II: Iterated Rules*, Intl. J. Intelligent Sys. **11**(8), 583-610, 1996.

12. Deng, X., Li, Y. and Deng, Y. *A Group Decision Making Method Based on Dempster-Shafer Theory of Evidence and IOWA Operator*, J. of Comp. Inf. Sys., **8**(9), 3929-3936, 2012.

13. Jin, L. and Xu, Y. *A Rule-Based Inference Method Using Dempster-Shafer Theory*, Knowledge Engineering and Management, Z. Wen and T. Li (eds), Adv. in Intel. Sys. and Comp., **278**, 61-72, 2014.

14. McKeever, S., Ye, J., Coyle, L. and Dobson, S. *Using Dempster-Shafer Theory of Evidence for Situation Awareness*, P. Barnaghi et al. (eds), EuroSSC 2009, LNCS **5741**, 149-162, 2009.

15. Mèrigo, J. M. and Casanovas, M. *Decision Making with Dempster-Shafer Theory Using Fuzzy Induced Aggregation Operators*, R. R. Yager et al. (eds): Recent Developments in the OWA Operators, Studies in Fuzziness and Soft Computing, **265**, 209-228, 2011.

16. Nusrat, E. and Yamada, K. *A Descriptive Decision-Making Model Under Uncertainty: Combination of Dempster-Shafer Theory and Prospect Theory*, Intl. J. of Uncert, Fuzz. and KBS, **21** (1), 79-102, 2013.

17. Bogler, P. L. *Shafer-Dempster Reasoning with Applications to Multisensor Target Identification Systems*, IEEE. Trans. on Sys., Man and Cyb., **17**(6), 968-977, 1987.

18. Dallil, A., Oussalah, M. and Ouldali, A. *Sensor Fusion and Target Tracking Using Evidential Data Association*, IEEE Sensors J., **13**(1), 285-293,2013.

19. Núñez, R. C., Samarakoon, B., Premaratne, K. and Murthi, M. N. *Hard and Soft Data Fusion for Joint Tracking and Classification/Intent-Detection*, Proc. 16th Intl. Conf. on Info. Fusion, Istanbul, Turkey, 661-668, 2013.

20. Stampouli, D., Brown, M. and Powell, G. *Fusion of Soft Information Using TBM*, In Proc. Intl. Conf. on Inf. Fusion, Edinburgh, Scotland, 1-8, 2010.

21. Wickramarathne, T. L. An Analytical Framework for Soft and Hard Data Fusion: A Dempster-Shafer Belief Theoretic Approach. PhD Thesis, University of Miami, 2012.

22. Wickramarathne, T. L., Premaratne, K., Murthi, M. N., Scheutz, M., Kubler, S. and Pravia, M. *Belief Theoretic Methods for Soft and Hard Data Fusion*, IEEE Intl. Conf. on Acoustics, Speech and Signal Proc. (ICASSP), 2388-2391, 2011.

23. Li, Y., Chen, J. and Feng, L. *Dealing with Uncertainty: A Survey of Theories and Practices*, IEEE Trans. on Knowledge and Data Eng., **25**(11), 2013.

24. Hadzagic, M., St-Hilaire, M-O., Valin, P. and Shahbazian, E. *Reliability and Relevance in the Thresholded Dempster-Shafer Algorithm for ESM Data Fusion*, Proc. 15th Intl. Conf. Info. Fusion, Singapore, 615-620, 2012.

25. Koks, D. and Challa, S. *An Introduction to Bayesian and Dempster Shafer Fusion*, DSTO Technical Report, DSTO-TR-1436, 2003.

26. Hensher, D. A. and Li, Z. *A Scoping Enquiry into the Potential Contribution of Subjective Probability Theory, Dempster-Shafer Theory and Possibility Theory in Accommodating Degrees of Belief in Traveller Behaviour Research*, Travel Behaviour and Society, **1**, 45-56, 2014.

27. Talavera, A., Aguasca, R., Gálvan, B. and Cacereño, A. *Application of Dempster-Shafer Theory for the Quantification and Propagation of the Uncertainty Caused by the Use of AIS Data*, Reliability Engineering and System Safely, **111**, 95-105, 2013.

28. Rekik, W., Le Hégart-Mascle, S., Reynaud, R., Kallel, A. and Hamida. A., *Dynamic estimation of the discernment frame in belief function theory*, Proc. 16th Intl. Conf. Info. Fusion, Istanbul, 2013, 1135-1142, 2013.

29. Wilson, N. *Algorithms for Demspter-Shafer Theory*, In Kohlas, J., Moral, S., (eds.) Algorithms for Uncertainty and Defeasible Reasoning, **5**, Handbook of Defeasible Reasoning, Kluwer Academic Publishers, 2000.

30. Oxenham, M. *The effect of finite set representations on the evaluation of Dempsters rule of combination*, In Proc. of Intl. Conf. on Info. Fusion, Cologne, Germany, 1-8, 2008.

31. Kennes, R. and Smets, Ph. *Computational Aspects of the M obius Transform*, Proc. 6th Conf. on Uncertainty in AI (UAI1990), 401-416, 1990.

32. Sentz, K. and Ferson, K. *Combination of Evidence in Dempster-Shafer Theory*, SAND Report, SAND 2002-0835, 2002.

33. Martin, A., *About Conflict in the Theory of Belief Functions*, Proc. Intl. Conf. on Belief Functions, Compiègne, France, 2012.

34. Smets, Ph. *The Nature of Unnormalized Beliefs Encountered in the Transferable Belief Model*, Proc. 8th Conf. on Uncertainty in AI, Dubois D., Wellmann M.P., DAmbrosio B. and Smets Ph. (eds), Morgan Kaufmann Publ., San Mateo, California, 292-297, 1992.

35. Zadeh, L. *A Mathematical Theory of Evidence* (book review), AI Magazine **5**(3), 81-83, 1984.

36. Schubert, J. *Conflict Management in Dempster-Shafer Theory Using the Degree of Falsity*, Intl. J. of Approx. Reasoning, **52**, 449-460, 2011.

37. Dezert, J. and Smarandache, F. Advances and Applications of DSmT for Information Fusion. Collected Works, **3**, Smarandache, F. and Dezert, J. (eds), Perfect Paperback, 92, 2009.

38. Dezert, J., Wang, P. and Tchamova, A. *On the Validity of Dempster-Shafer Theory*, Proc. Fusion 2012 - 15th Intl. Conf. on Inf. Fusion, 655-660, 2012.

39. Tchamova, A. and Dezert, J. *On the Behavior of Dempster's Rule of Combination and the Foundations of Dempster-Shafer Theory*, Proc. 6th IEEE Intl. Conf. in Intel. Sys., 108-113, 2012.

40. Smets, P. *Data Fusion in the Transferable Belief Model*, Proc. 3rd Intl. Conf. Inf. Fusion, Paris, France, 21-33, 2000.

41. Delmotte, F. and Smets, Ph. *Target Identification Based on the Transferable Belief Model Interpretation of Dempster-Shafer Model. Part I: Methodology. Part 2: Application*, IRIDIA Technical Report, 2001.

42. Delmotte, F. and Smets, Ph. *Target Identification Based on the Transferable Belief Model Interpretation of Dempster-Shafer Model*, IEEE Trans. Syst., Man, Cybern. A:34, 457-471, 2004.

43. Ristic, B. and Smets, Ph. *Target identification using belief functions and implication rules*, IEEE Trans. on Aero. and Elec. Sys., **41**(3), 1097-1103, 2005.

44. Ristic, B. and Smets, Ph. *Kalman filter and Joint Tracking and Classification in the TBM Framework*, In Proceedings of the Seventh International Conference on Information Fusion , **1**, 46-53, 2004.

45. Ristic, B. and Smets, Ph. *Belief function theory on the continuous space with an application to model based classification*, Proc. Info. Proc. and Managmenet of Uncert (IPMU), 1119-1126, 2004.

46. Smets, Ph. and Ristic, B. *Kalman Filter and Joint Tracking and classification based on Belief Functions in the TBM Framework*, Information Fusion, **8**(1), 16-27, 2007.

47. Powell, G., Marshall, D., Smets, P., Ristic, B., and Maskell, S. *Joint Tracking and Classification of Airborne Objects Using Particle Filters and the Continuous Transferable Belief Model*, In Proceedings of the 9th International Conference on Information Fusion, 1-8, 2006.

48. Ristic, B., and Smets, Ph. *Target classification approach based on the belief function theory*, IEEE Transactions on Aerospace and Electronic Systems, **41**(2), 574-583, 2005.

49. Yager, R. *On the Dempster-Shafer Framework and New Combination Rules*, Inf. Sciences, **41**, 93-137, 1987.

50. Inagaki, T. *Interdependence Between Safety-Control Policy and Multiple-Sensor Schemes Via Dempster-Shafer Theory*, IEEE Trans. on Reliability, **40**(2), 182-188, 1991.

51. Dubois, D. and Prade, H. *A Set-Theoretic View on Belief Functions: Logical Operations and Approximation by Fuzzy Sets*, Intl. J. Gen. Sys., **12**, 193-226, 1986.

52. Zhang, L. *Representations, Independence, and Combination of Evidence in the Demspter-Shafer Theory*, Advances in the Dempster-Shafer Theory of Evidence, 51-69, 1994.

53. Ferson, S. and Kreinovich, V. *Representation, Propagation, and Aggregation of Uncertainty in Risk Analysis - from Traditional Probabilistic Techniques to More General, More Realistic Approaches: A Survey*, University of Texas, El Paso Report, UTEP-CS-01-33, 2001.

54. Jøsang, A., Diaz, J. and Rifqi, M. *Cumulative and Averaging Fusion of Beliefs*, Inf. Fusion, **11**(2), 192-200, 2010.

55. Denœux, T. *The Cautious Rule of Combination for Belief Functions and Some Extensions*, Proc. 9th Intl. Conf. on Inf. Fusion, Florence, Italy, 2006.

56. Florea, M. C., Dezert J., Valin, P., Smarandache, F., Jousselme, A-L. *Adaptive Combination Rule and Proportional Conflict Redistribution Rule for Information Fusion*, In Proc. of COGIS 06, Cogn. Sys. with Interactive Sensors, Paris, 2006.

57. Dubois, D. and Prade, H. *Representation and combination of uncertainty with belief functions and possibility measures*, **4**(3), Computational Intelligence, 244-264, 2007.

58. Smets, Ph. *Analyzing the Combination of Conflicting Belief Functions*, Inf. Fusion, **8**(4), 387-412, 2007.

59. Nguyen, V. and Docking. M. *Infux: A Tool and Framework for Reasoning under Uncertainty*, Draft DSTO Technical Report, 2014.

60. Oxenham, M., Challa, S. and Morelande, M. *Fusion of Disparate Identity Estimates for Shared Situation Awareness in a Network-Centric Environment*, Inf. Fusion, **7**, 395-417, 2006.

61. Sudano, J. J. and Martin, L. *Yet Another Paradigm Illustrating Evidence Fusion (YAPIEF)*, Fusion 2006, 1-7, 2006.

62. Cuzzolin, F. *The Intersection Probability and its Properties*, Sossai, C. and Chemello G., (Eds.): Proceedings of the 10th European Conference on Symbolic and Quantitative Approaches to Reasoning with Uncertainty (ECSQARU), Verona, Italy, pp. 287298, 2009.

63. Klir, G. J. Uncertainty and Information: Foundations of Generalized Information Theory, 226-228, 2006.

64. Cobb, B. R. and Shenoy, P. P. *On the Plausibility Transformation Method for Translating Belief Function Models to Probability Models*, Intl. J. Approx. Reasoning, **41**(3), 314-330, 2006.

DISTRIBUTION LIST

A Discussion of Dempster-Shafer Theory and its Application to Identification Fusion

E. El-Mahassni and K. White

AUSTRALIA

DEFENCE ORGANISATION

Task Sponsor

Program Manager Joint Strike Fighter: AVM Chris Deeble	1
Director General JSF Acquisition and Sustainment: AIRCDRE Terrence Saunder	1
CASG JSF Software Lead: FLTLT Matthew Gibbons	1
A/Director S&T JSF IPT: Mr. Rick Dykstra	1
Program Leader Air Combat Capability: Mr. Khan Sharp	1
Chief of National Security and ISR Division: Dr. Tony Lindsay	1
Research Leader Information Integration Branch: Dr. John Percival	1
Group Leader Data and Information Fusion Group: Prof. Neil Gordon	1
JSF Fusion S&T Lead: Dr. Krüger White	1
Author: Dr. Edwin El-Mahassni	1

Page classification: UNCLASSIFIED

DEFENCE SCIENCE AND TECHNOLOGY GROUP DOCUMENT CONTROL DATA	1. DLM/CAVEAT (OF DOCUMENT)

2. TITLE	3. SECURITY CLASSIFICATION (FOR UNCLASSIFIED REPORTS THAT ARE LIMITED RELEASE USE (L) NEXT TO DOCUMENT CLASSIFICATION)
A Discussion of Dempster-Shafer Theory and its Application to Identification Fusion	Document (U) Title (U) Abstract (U)
4. AUTHORS	5. CORPORATE AUTHOR
E. El-Mahassni and K. White	Defence Science and Technology Group PO Box 1500 Edinburgh, South Australia 5111, Australia

6a. DST GROUP NUMBER	6b. AR NUMBER	6c. TYPE OF REPORT	7. DOCUMENT DATE
DST-Group–TN–1443	016-371	Technical Note	August, 2015

8. FILE NUMBER	9. TASK NUMBER	10. TASK SPONSOR	11. No. OF PAGES	12. No. OF REFS
			24	64

13. DST Group Publications Repository	14. RELEASE AUTHORITY
`http://dspace.dsto.defence.gov.au/dspace/`	Chief, National Security and ISR Division

15. SECONDARY RELEASE STATEMENT OF THIS DOCUMENT

Approved for Public Release

OVERSEAS ENQUIRIES OUTSIDE STATED LIMITATIONS SHOULD BE REFERRED THROUGH DOCUMENT EXCHANGE, PO BOX 1500, EDINBURGH, SOUTH AUSTRALIA 5111

16. DELIBERATE ANNOUNCEMENT

No Limitations

17. CITATION IN OTHER DOCUMENTS

No Limitations

18. DST GROUP RESEARCH LIBRARY THESAURUS

Sensor Fusion, Data Fusion, Target Classification, Decision Theory

19. ABSTRACT

This paper outlines some of the basics of Dempster-Shafer Theory, which is a mathematical theory for combining evidence from different sources to obtain a degree of belief in a proposition. In particular, different combination rules available within the context of Identification Fusion and the assumptions and implications of each of those rules are outlined and investigated. However, the belief function arising from combining evidence under Dempster-Shafer Theory is often insufficient to support decision-making, and a transformation from the belief function to a probability distribution is required. Several different transformations and illustrative examples implementing Dempster-Shafer Theory for Identification Fusion are provided. The results and some possible directions for future work are discussed.

Page classification: UNCLASSIFIED